METHODS AND SYSTEMS FOR READ-ALOUD BOOKS

Related Application Data

This application is a continuation-in-part of copending application 09/571,422, filed May 15, 2000, which claims priority to provisional application 60/151,586, filed August 30, 1999.

Field of the Invention

The present invention relates to entertainment devices and methods, and is particularly illustrated in the context of read-aloud children's books.

Background and Summary of the Invention

Children learn the mechanics of turning book pages at an early age. They enjoy looking at pictures on pages, and hearing any accompanying text read-aloud. Many children have favorite books that they like to hear read over and over again. The association of seeing pictures and repeatedly hearing the words is an excellent mechanism for learning to read and learning to enjoy books.

In accordance with one embodiment of the present invention, these childhood reading experiences are complemented through use of machine-readable data in books that permit children to hear the text of a book even when an adult is not available to assist them.

In one particular arrangement, the pages of a book are marked in a generally human-imperceptible manner using digital watermark technology. For example, a story illustration (or the background of a page of text) can be subtly changed, in local luminance detail, to encode a 32 bit number (a page identifier). When the child presents the page to a web cam or the like, associated image data is sensed and processed to recover the 32 bit identifier. This identifier can then be used (locally, or by transmission to a remote data repository) to index multimedia data that is then presented to the child, as read-aloud spoken text.

The foregoing and additional features and advantages of the invention will be more readily apparent from the following detailed description, which proceeds with reference to the accompanying drawings.

Brief Description of the Drawings

Fig. 1 is a block diagram depicting one embodiment of the present invention.

Detailed Description

Referring to Fig. 1, a system 10 according to one embodiment of the invention includes a toy 12 and a computer device 14. The toy (e.g., a doll or stuffed animal) includes, concealed within it, an image sensor 16 and a speaker 18. (The image sensor may view out of one of the toy's eyes.) The computer can be as complex or as simple as system constraints dictate. Typically, the computer includes a processor and a memory. A display 20 may be integrally provided as part of the computer device, or it may be separate.

In use, a child presents a page of a book 22 to the image sensor within the toy 12. Image data acquired by the sensor is passed to the computer device, and analyzed to decode the machine readable page identifier data therein. This identifier is then used to index a data store that includes, e.g., text or digitized speech data corresponding to the presented page. This data can be local (e.g., stored on a hard disk within the computer device 14) or remote. In the later case, the computer device can use known means, such as the internet, to link to a remote repository at which the speech data is stored. This data is then transferred back to the local system, where it is rendered for the child. The child can then show the next page to the toy and repeat the process.

The foregoing description describes an elementary system. A great number of enhancements can be made.

For example, speech is not the only response that can be triggered by display of a book page to the sensor. Essentially any other response or action can also be triggered. One is to present supplementary image or video information on the display 20, such as

alternate illustrations. Another is to retrieve line artwork related to the story that can be printed by the child and colored. Etc.

If the data repository is remote, the data latency between sending a request for speech data, and receiving the requested speech data back, may introduce an objectionable delay. Accordingly, the system can be arranged to pre-fetch anticipated data. Thus, for example, if a child presents page 1 of The Cat in the Hat to the image sensor, and the computer queries the remote data repository for the page 1 speech data, it can thereafter immediately solicit the data corresponding to pages 2, 3, 4, etc., and cache such data locally for their anticipated next use.

The narration accompanying a book can be customized with the reader's name, street name, hometown, friends' names, etc., as appropriate. Thus, the hero of a book can be given the reader's first name, live in the same town, etc. (Such data can be stored locally, e.g., in an XML data file indicating <HERO>="David"; <TOWN>="Portland," etc.) The narration text can be flagged at places with corresponding XML tags. When the narration text comes to a <HERO> tag, it can check to see whether customized <HERO> data is available and, if so, whether it has been enabled for use. In such case, the name "David" will be substituted. Else, the usual (default) hero name from the narration data will be employed.

Yet another enhancement is to let the reader chose how the plot unfolds. Such "chose your own adventure" books are now popular with older juvenile readers, and specify to what page the reader should turn, depending on the desired story action. Thus, the bottom of a page may indicate "Turn to page 137 if you want Nancy Drew to climb down into the cave; turn to page 216 if you want her to avoid the cave." Through such branching, the reader can explore many "what if" plot scenarios within a single book.

Such functionality can also be achieved in accordance with the present invention. For example, the rendered speech from one page can instruct the child, "If you want the Cat in the Hat to open mother's closet door, move the book up and down. If you want the Cat in the Hat to stay out of mother's closet, move the book from side to side." The computer can detect such gestures from the image data, and invoke different actions (e.g., tell different story lines, or present different video accompaniment) accordingly.

Book pages can also be used as indices to DVD movies and the like. Consider the popular Harry Potter books, for which corresponding movies are being produced. If a child wants to view the movie excerpt where Harry fights the dragon, that page of the book can be shown to an image sensor. The computer decodes a watermark from that page and instructs a DVD player associated with the system to advance to a corresponding frame number and begin playing.

In such systems, different book pages are watermarked with different identifiers. (Each page may have a unique identifier, or all pages in a chapter may have a common identifier, etc.) These identifiers may be literal frame numbers of the corresponding video production, but more generally are not. Instead, the mapping of page identifiers to frame numbers is typically performed by a data structure, such as a database, that has a record for each identifier, containing the corresponding page number and, perhaps, other information. The output from the data structure is provided to the control processor in the DVD player, and causes it to advance to the specified video location and begin playing.

The parent application details a great variety of systems in which physical objects are encoded with digital data, sensed with a web cam or the like, and then used to invoke some computer-controlled action. Illustrative systems and infrastructures associated with such technologies are detailed in that application and so are not belabored here.

Likewise, a great variety of particular watermarking techniques are known to artisans in the field, and are not belabored here. One particular technique is shown in application 09/503,881, filed February 14, 2000.

To provide a comprehensive disclosure without unduly lengthening this specification, the above-cited patent applications are incorporated herein by reference.

Having described and illustrated the principles of the invention with reference to illustrative embodiments, it should be recognized that the invention is not so limited.

For example, while the preferred embodiment employed digital watermarking to mark pages, other machine-readable data representations can be used instead (e.g., bar codes, glyphs, RF IDs, mag stripes, smart card technology, etc.).

More generally, the invention is not limited to use with children's books. For example, the same principles can be employed to provide read-aloud functionality with adult reading materials, e.g., for the blind or visually impaired. Moreover, the invention is not limited to books, but can be used in conjunction with any object for which supplementary audio or video would enhance a user experience (e.g., baseball cards that link to videos of famous plays; comic books that link to previews of the next issue, etc.)

The principles of the invention can be used in contexts other than reading-aloud. For example, showing a book page (or book cover) to the system can initiate playback of mood music to accompany the child's independent reading (silent, or aloud) of the story.

Showing the page to the image sensor can alternatively link to an interactive web page. For example, a watermark in the book Yertle the Turtle (or on a front/back cover) can lead to a web page focused on turtles; a watermark in The Lorax can link to an environmental web page tailored for children. Etc.

Likewise, the watermark can link to an interactive internet game, or other immersive experience. (E.g., Nancy Drew finds herself in a cave. The reader, together with hundreds of other children who are reading the same story at the same time, hunt for clues to solve the mystery – either independently or collaboratively. Prizes may be awarded to readers who solve the puzzle, such as a five dollar coupon for Amazon.com. Through choices the reader makes, an interactive web experience may lead to a different story outcome than the printed book.)

The present technology can be used with media other than books. For example, a child may receive a watermarked birthday card from Grandma. When the card is shown to the plush toy, the system could place a telephone call to Grandma – allowing her to convey personal birthday greetings to the child.

The reference to computer devices in the foregoing discussion should not be taken as limiting applicability of the invention to any particular form of hardware (e.g., desktop computers). For example, a cell phone equipped with an optical sensor, a headset, and internet connectivity can be employed to sense book pages, and render corresponding audio – a boon for families on long car trips. Set top boxes can also provide such functionality, and make use of the audio and display capabilities of an associated

television receiver. Popular gaming platforms (e.g., PlayStation 2) can be adapted for such use. A plush toy incorporating the camera can be self-contained with a built-in processor, and be linked to a display device (e.g., a television) by wireless (e.g., Bluetooth). Etc.

The incorporation of the web cam and the speaker within a plush toy is suitable in some embodiments, but in other embodiments these devices can be incorporated into different objects (e.g., the housing of a computer display), or used independently.

While reference was made to DVD videos that are indexed by data encoded on book pages, the same principles can be employed in other contexts. For example, instead of indexing a video stored on DVD, the video can be obtained from a remote source, such as a cable TV system or streaming internet video, and rendered starting from a point determined by the book page.

A number of different text-to-speech technologies are known, and so are not belabored here. But this aspect of the system is subject to many variations. For example, the voice with which a book is read can be the child's parent. The parent can record a read-aloud session, with each page of reading being stored in a different data file (or database record). Playback of these files/records can then be triggered by detection of watermarks from the corresponding book pages. Similarly, the child's own voice can be employed in the same manner.

While the foregoing approach requires each book to be read-through once in order to record the voice, other approaches do not have this requirement. Instead, the speech can be synthesized from component utterances (e.g., common words and phonemes) that are patched together to create arbitrary speech. (Such systems require training, e.g., by the speaker reading a prescribed passage of text from which the component words and sounds are recognized, parsed, and stored for later use.) The component parts can be stored locally and assembled by the computer 14. Or they can be stored remotely (e.g., accessed over the internet), and assembled either by the computer 14 or a remote computer. By such arrangements, a child can be read a story by a familiar voice, even if that person has never before seen the book.

A variety of celebrity and family voices can be available for read-aloud purposes, using the foregoing or other approaches. Thus, for example, a child can select to have a story read by Barney, Sesame Street's Cookie Monster, mom, dad, a favorite uncle, etc. Images or video associated with the reader can be presented on the display screen during the reading.

The above-described substitution of the child's name, hometown, etc., in the story line is one variant of "Mad Libs," in which arbitrary words are substituted into a text. Other such Mad Lib variants can be employed. For example, after a child has become familiar with a story (e.g., a poem from the Monster Motel book), the stuffed animal (i.e., the computer, through the speaker) can ask the child if she would like to play a game. The animal could then explain what an adjective is – a word that describes something, such as "blue" or "funny" – and ask the child to propose such a word. Likewise the animal could ask the child for one or more nouns, verbs, and adverbs. After recording a suitable collection of words, the system could read the poem, substituting the child's words for key words in the poem.

Reference was made, above, to the system being responsive to gestures made by the child with the book, triggering different responses. This area, too, is subject to a great number of variants. For example, one gesture could cause reading of the page to start over. Another could speed up the voice. Another could select among different reading voices. Another could initiate internet-based functionality (as described above). Another could change the reading volume, etc. (Many gestures, and methods of their detection, are detailed in the parent application.)

It should be recognized that the particular combinations of elements and features in the above-detailed embodiments are exemplary only; the interchanging and substitution of these teachings with other teachings in this and the incorporated-by-reference patents/applications are also contemplated.

In view of the wide variety of embodiments to which the principles and features discussed above can be applied, it should be apparent that the detailed embodiments are illustrative only and should not be taken as limiting the scope of the invention. Rather,

we claim as our invention all such modifications as may come within the scope and spirit of the following claims and equivalents thereof.